When Social Learning meets Machine Learning: Virtues and Vices of the digital economy

Fabrizio Germano [1] and Stefano Lovo [2]

[1] Universitat Pompeu Fabra

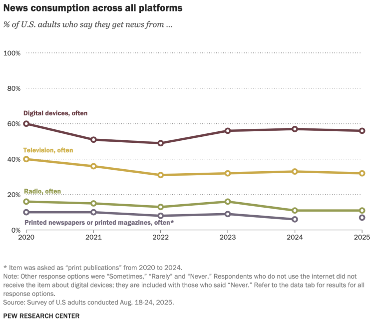
[2] HEC, Paris

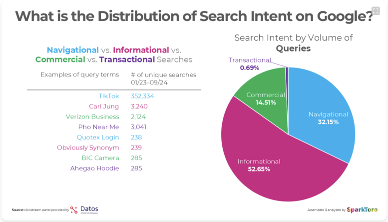
March 17, 2026

Motivation

How people get informed

- The internet has become the main platform people get news from.
- Out of the 9.5 million Google searches per minute, 53% are informational and about 15% are commercial.





Motivation

The main source of revenue in the internet economy

- Internet platforms such as search engines and social media generate a large share of their revenue through **personalized online advertising** that makes firms visible to the users most likely to buy their products.
- Personalized online advertising is enabled by machine-learning algorithms that, through 'cookies', personalize the content displayed to users.
- Algorithms are optimized to maximize the time users spend on these platforms. More time spent on the platform:
  - better knowledge of users' tastes
  - more exposure to the 'right' ads
  - greater effectiveness of ads
  - higher fees the platform can charge to advertisers

Research questions

Consider a platform whose algorithm is parameterized so as to maximize the expected time users spend on the platform.

How does this affect:

1. Matching between firms and their potential customres among the SE users
2. Whether, in the process of being matched to contents, users
   - obtain meaningful and objective information
   - converge to a consensus on essential issues of public interest, across different informational environments.
3. In the absence of consensus, what drives the divergence and how strong is it.

Model

- A mass 1 of atomistic users $i \in [0, 1]$
- A mass 1 of atomistic firms $j \in [0, 1]$, each with a website of content c_j
- A search engine (SE)
- Infinitely many rounds $\tau = -1, 0, 1, 2, \ldots$
- At $\tau = -1$, Nature chooses a state $\omega \in \{0, 1\}$, with $\Pr(\omega = 1) = \pi = 0.5$
- In each round $\tau > 0$:
    1. Each user i submits an internet query to the SE
    2. The SE matches each user i to a website j
    3. Each user i observes $c_{\tau,i}$, the content of the website he has matched with, updates his/her belief about ω, and chooses how much time $t_{\tau,i} \in [0, 1]$ to spend consulting website j

Internet content

State dependent and exogenous

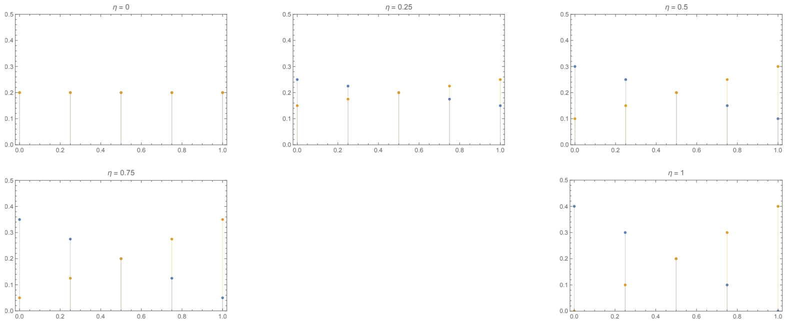
- $X \subset [0,1]$: Finite countable set of all content present on the web.
- for any $c \in X$, let $g_\omega(c) \in [0,1] :=$ (exogenous) fraction of firms with website content $c_j = c$ if the state of Nature is ω.
- Posterior from unbiased sampling:

$$Pr(\omega = 1|c) = P(c, \pi) := \frac{\pi g_1(c)}{\pi g_1(c) + (1-\pi) g_0(c)}$$

Assumptions g_ω satisfies:

- Monotonicity: $\frac{\partial P(c,\pi)}{\partial c} > 0$
- Symmetry: $P(c, \pi) = 1 - P(1-c, \pi)$
- Web-informativeness parameter $\eta \in [0,1]$: $\frac{\partial^2 P(c,\pi)}{\partial c \partial \eta} > 0$

Distribution of content



The figure reports the values of $g_0(c)$, in blue, and $g_1(c)$, in orange, for $c \in X = \{0, 0.25, 0.5, 0.75, 1\}$. Each panel corresponds to a different value of $\eta \in \{0, 0.25, 0.5, 0.75, 1\}$.

Web-informativeness parameter $\eta \in [0,1]$

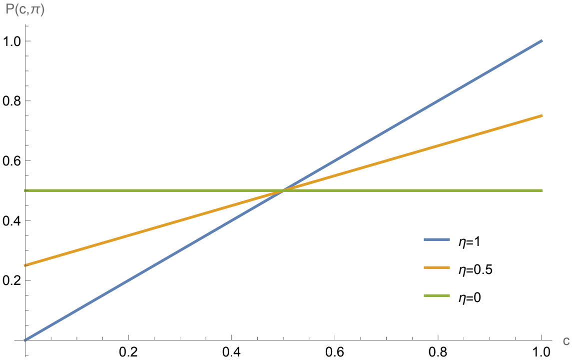


Figure: Posterior belief for prior $\pi = 1/2$ as a function of content c for different values of η.

User preferences

- Each user is characterized by a privately known idiosyncratic type $x_i \in X$, with $x_i \perp \omega$.
- Enjoyment of user i from the marginal time spent consulting content c if the state is ω:

$$V(x_i, c, \omega) := 1 - \rho(x_i - c)^2$$

  - User prefers content c closer to his/her idiosyncratic type x_i.

User preferences

- Each user is characterized by a privately known idiosyncratic type $x_i \in X$, with $x_i \perp \omega$.
- Enjoyment of user i from the marginal time spent consulting content c if the state is ω:

$$V(x_i, c, \omega) := 1 - \rho(x_i - c)^2 - (1 - \rho)(\omega - c)^2$$

  - User prefers content c closer to his/her idiosyncratic type x_i.
  - User prefers content c closer to the true state of Nature ω.

User preferences

- Each user is characterized by a privately known idiosyncratic type $x_i \in X$, with $x_i \perp \omega$.
- Enjoyment of user i from the marginal time spent consulting content c if the state is ω:

$$V(x_i, c, \omega) := 1 - \rho(x_i - c)^2 - (1 - \rho)(\omega - c)^2$$

  - User prefers content c closer to his/her idiosyncratic type x_i.
  - User prefers content c closer to the true state of Nature ω.
  - $\rho \in [0, 1]$ is the query-specific parameter measuring the extent to which the query is commercial (private value) versus informational (common value).

User consulting time

- Each user privately knows his/her type $x_i \perp \omega$ and does not observe ω.
- In round τ, user i is matched with website of content c, chooses consulting time t that solves:

$$\max_t E\left[V(x_i, c, \tilde{\omega})t - \frac{t^2}{2} \middle| x_i, c_{\tau,i}, h_{\tau,i} \right]$$

- Let $\pi_{\tau,i}$ denote user i's belief in round τ, and $\omega = 1$, then consulting time is

$$t^\star(x_i, c, \pi_{\tau,i}) := \\ 1 - \rho(c - x_i)^2 - (1-\rho)(\pi_{\tau,i}(c-1)^2 + (1-\pi_{\tau,i})c^2)$$

Two key exogenous parameters ρ and η to encompass most internet queries

$(\eta, \rho) \in [0,1] \times [0,1]$	low η	high η
low ρ	contested informational queries example: Is red wine good or bad for health?	consensual information queries example: capital of Nebraska
high ρ	commercial queries on new products example: best AI for my use	commercial queries on known products example: best jeans for me

Some preliminary remarks

- There is a continuum of firms whose content distribution is state dependent $\Rightarrow \omega$ is fully learnable if $\eta > 0$.

- There is a continuum of rational users with i) common prior belief and ii) whose type x_i is not correlated with $\omega \Rightarrow$ no specific reason for type x_i to affect a user posterior belief

First best with omniscient benevolent planner willing to maximize ex-post utility

Match user i with the website content $c \in X$ closest to

$$\hat{c}(x_, \omega) = \rho x_i + (1 - \rho)\omega$$

From observing whether c_i equals $\hat{c}(x_i, 0)$ or $\hat{c}(x_i, 1)$ the user learns the state and is given his/her personalized best possible content.

Private histories at the beginning of round τ

- No one directly observes ω
- Users i: x_i and browsing history: $h_{\tau,i} := \{c_{\tau',i}, t_{\tau',i}\}_{\tau'<\tau}$

  where:
  $c_{\tau',i}$ is the content resulting from i's query in round τ';
  $t_{\tau',i}$ is user i's consulting time of $c_{\tau',i}$;

- Search Engine $h_{\tau,SE}$:
  - website contents : $c_j, \forall j \in [0,1]$
  - Users browsing histories: $\{c_{\tau',i}, t_{\tau',i}\}_{\tau'<\tau}, \forall i \in [0,1]$

Search engine q-learning algorithm

Users-content matching

At any given τ, the SE matches each user i to a website j using a **q-learning algorithm** that accesses h_τ^{SE}.

Search engine q-learning algorithm

Users-content matching

At any given τ, the SE matches each user i to a website j using a **q-learning algorithm** that accesses h_τ^{SE}.

- Q-matrix: $Q_{\tau,i}(c)$ is the value that, in round τ, the SE algorithm associates with matching user i to a website with content c.
- Round τ matching:
  - With probability $1 - e^{-\beta\tau}$, the SE **exploits**: it matches user i to *greedy content*
  $$c_{\tau,i} = \hat{c}_{\tau,i} := \arg\max_{c \in X} Q_{\tau,i}(c)$$
  - With probability $e^{-\beta\tau}$, the SE **explores** and matches i with a random firm $j \in [0,1]$. Under uniform random selection of j,
  $$\Pr(c_{\tau,i} = c|\omega) = g_\omega(c)$$

Search engine q-learning algorithm

Q-matrix updating

- Q-matrix initialization:

$$Q_{1,i}(c) = 0, \forall i, c$$

- Q-matrix evolution:

$$Q_{\tau+1,i}(c) = \alpha \bar{t}_{\tau,i}(c) + (1-\alpha) Q_{\tau,i}(c)$$

where $\bar{t}_{\tau,i}(c)$ is the average consulting time in round τ of all users who:

i) have the same browsing history $h_{\tau,i'}$ as user i's,
and
ii) were displayed content c in round τ.

In words:

- Because the number of types x_i is finite and the probability of a user being exposed to content c is at least $e^{-\beta\tau} g_\omega(c) > 0$, for any $c \in X$ in every round τ, there is a positive mass of users of your same browsing history before τ who have been exposed to c in round τ.
- The SE uses the consulting time of users who had your same browsing history to update the value of your Q-matrix for all contents, including those content you have not been exposed to in this round.

Search engine q-learning algorithm

Algorithm parametrization

- α:= weight of last observed consulting time in the Q-matrix values updating.
- β:= decay rate of exploring probability.

Algorithm parameters α and β are chosen by the SE owner so as to **maximize the time users on average spend consulting websites**.

Users' beliefs Bayesian updating

- Each user understands the functioning of the SE algorithm and correctly anticipate α and β parametrization, and from $h_{\tau,i}$ deduces $\hat{c}_{\tau,i}$.
- Whenever $c_{\tau,i} \neq \hat{c}_{\tau,i}$, user i knows this comes from SE exploring: unbiased internet sampling.
- Whenever $c_{\tau,i} = \hat{c}_{\tau,i}$, user i is not certain this comes from unbiased sampling or from SE exploiting.

$$\pi_{\tau,i} = Pr(\omega = 1 | c_{\tau,i}, h_{\tau,i}) = \begin{cases} P(c_{\tau,i}, \pi_{\tau-1,i}) \text{ if } c_{\tau,i} \neq \hat{c}_{\tau,i} \\ \hat{P}(c_{\tau,i}, \pi_{\tau-1,i}, \tau) \text{ if } c_{\tau,i} = \hat{c}_{\tau,i} \end{cases}$$

where

$$\hat{P}(c, \pi, \tau) := \frac{\pi(e^{-\beta\tau} g_1(c) + 1 - e^{-\beta\tau})}{e^{-\beta\tau}(\pi(g_1(c) + (1-\pi)g_0(x)) + 1 - e^{-\beta\tau}}$$

$$\begin{aligned} P\left(\frac{1}{2}, \pi\right) = \hat{P}\left(\frac{1}{2}, \pi, \tau\right) &= \pi \\ \frac{\partial P}{\partial c} &> 0 \\ \frac{\partial \hat{P}}{\partial c} &> 0 \\ |\hat{P}(c, \pi, \tau) - \pi| &< |P(c, \pi) - \pi| \end{aligned}$$

Content personalization and beliefs

Users are rational and start from a common prior belief, but posterior beliefs can diverge because of:

- Differences in queries outcomes: Each user is overexposed to the greedy content the SE associates to him/her
- Differences in interpretation: discounts information content of $c_{\tau,i}$ if $\hat{c}_{\tau,i} = \hat{c}_{\tau,i}$.

Effect of content personalization

- Heterogeneity in posterior beliefs
- 'Rabbit-hole' long-term beliefs:
  For any small $\varepsilon > 0$ there is $\underline{\tau}$ such that for every user i and any given level of belief $\pi_{\underline{\tau},i} \in [0,1]$,

$$\Pr\left(\frac{\pi_{\tau',i}}{\pi_{\underline{\tau},i}} \in [1-\varepsilon, 1+\varepsilon], \forall \tau' \geq \underline{\tau}\right) > 1 - \varepsilon \tag{1}$$

All these effects are stronger the higher the probability the SE is exploiting (higher β).

SE parametrization of α and β to maximize users consulting time.

To maximize consulting time exploration-decay-rate β increases in ρ and in η.

$\eta \setminus \rho$	0.0	0.2	0.4	0.6	0.8	1.0
0.25	(0.80,0.02)	(0.80,0.02)	(0.80,0.02)	(0.80,0.02)	(0.80,0.04)	(0.50,2)
0.50	(0.80,0.04)	(0.80,0.04)	(0.80,0.04)	(0.80,0.04)	(0.80,0.06)	(0.05,2)
0.75	(0.80,0.04)	(0.80,0.04)	(0.80,0.06)	(0.80,0.06)	(0.50,0.10)	(0.01,2)
1.00	(0.01,0.10)	(0.50,0.06)	(0.50,0.10)	(0.50,0.10)	(0.50,0.20)	(0.00,2)

Table: Average Consulting time maximizing Parameters (α, β), given $X = \{0, 0.25, 0.5, 0.25, 1\}$.

SE pretty good for commercial queries not so good for contested information queries

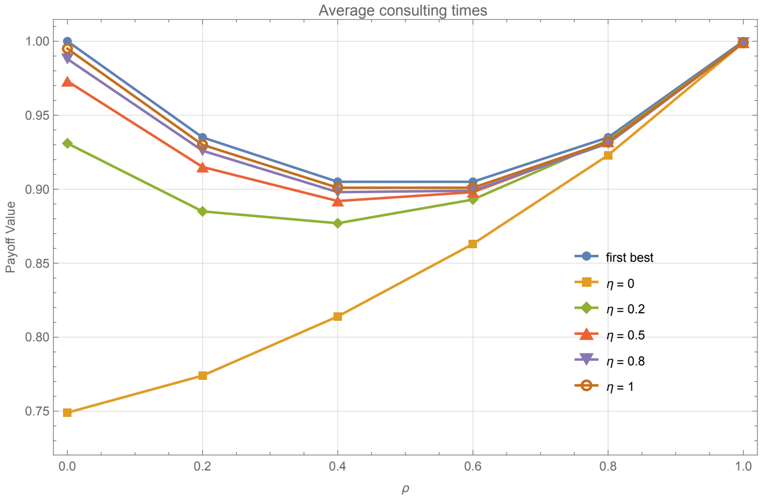
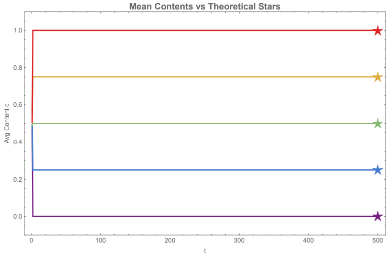
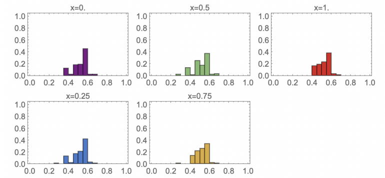


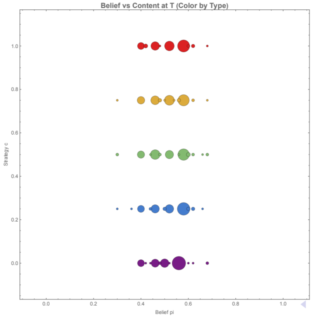
Figure: Users average browsing time as function of the commercial nature of the query ρ for different level of web informativeness η, given the average browsing time maximization choice of (α, β) for any given (η, ρ).

Pure commercial queries $X = \{0, 0.25, 0.5, 0.75, 1\}$

$\{\rho, \eta\} = \{1, 0.25\} \Rightarrow \beta = 2$

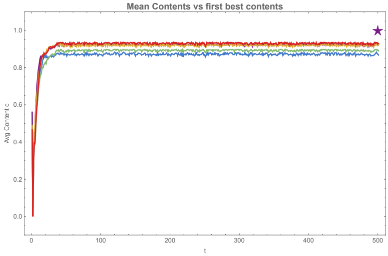


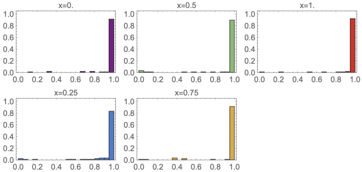


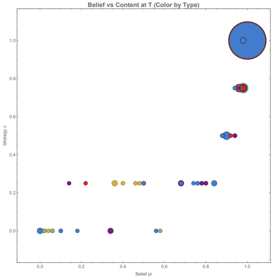


Pure information consensual queries $X = \{0, 0.25, 0.5, 0.75, 1\}$

$\{\rho, \eta\} = \{0, 1\} \Rightarrow \beta = 0.1$





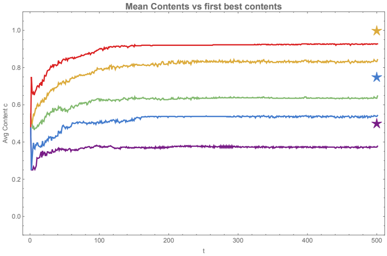


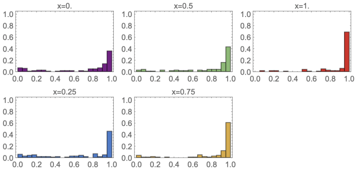
5% of users with $\pi_{500,i} < \frac{1}{2}$

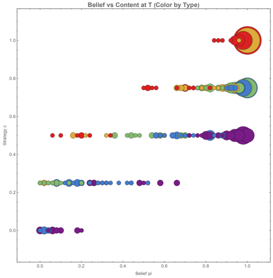
Figure: Simulations for $\omega = 1$

Information contested queries

$\{\rho, \eta\} = \{0.4, 0.25\} \Rightarrow \beta = 0.02$





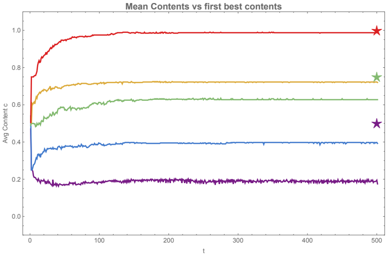


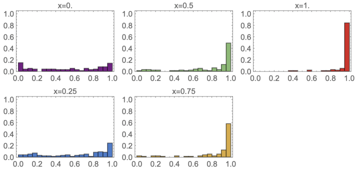
22.6% of users with $\pi_{500,i} < \frac{1}{2}$

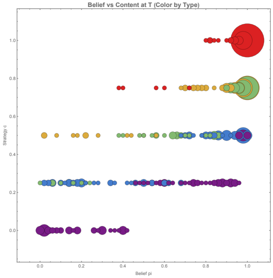
Figure: Simulations for $\omega = 1$

Commercial queries on new products

$\{\rho, \eta\} = \{0.6, 0.25\} \Rightarrow \beta = 0.02$







22.3% of users with $\pi_{500,i} < \frac{1}{2}$

Figure: Simulations for $\omega = 1$

Equilibrium with rational expectation users:

- SE algorithm is very effective for
  - Pure commercial queries
  - Pure informational consensual queries
- For hybrid queries platform performance is poorer for users whose types are opposite to the state.
  - Contents of extreme types opposite to the state are overly biased toward the type.
  - Beliefs of user with extreme type opposite to the state are more dispersed (sometimes completely wrong) compared average users.

Naive belief updating

Users do not fully grasp how the algorithm functions and take query outcomes to be an unbiased sample from the internet:

$$\pi_{\tau,i}(c_{\tau,i}) = P(c_{\tau,i}, \pi_{\tau-1,i})$$

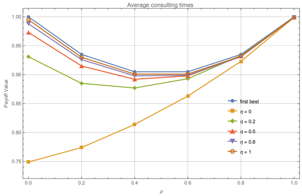
Users are overexposed to greedy content, but do not account for this over-exposition when updating their beliefs

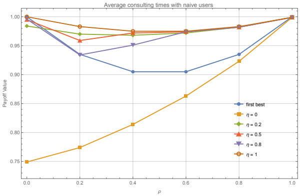
$$\Rightarrow$$

non-neutral greedy content may lead to extreme beliefs.

SE parametrization if facing naive users

$\eta \setminus \rho$	0.0	0.2	0.4	0.6	0.8	1.0
0.00	(0.80,2)	(0.80,2)	(0.80,2)	(0.80,2)	(0.80,2)	(0.05,2)
0.25	(0.05,2)	(0.05,2)	(0.05,2)	(0.10,2)	(0.80,2)	(0.50,2)
0.50	(0.05,2)	(0.10,2)	(0.10,2)	(0.10,2)	(0.80,2)	(0.05,2)
0.75	(0.10,2)	(0.10,0.20)	(0.50,2)	(0.10,2)	(0.50,2)	(0.01,2)
1.00	(0.00,2)	(0.50,2)	(0.10,2)	(0.10,2)	(0.10,2)	(0.00,2)

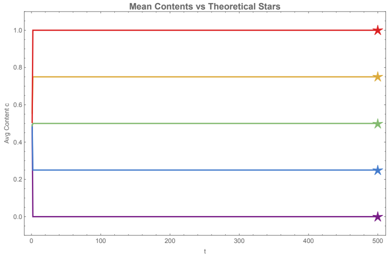


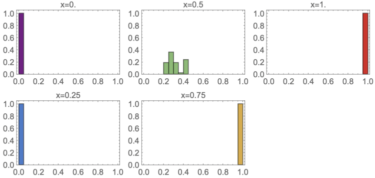


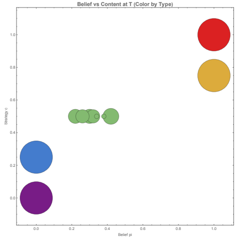
Pure commercial queries with naive users

$X = \{0, 0.25, 0.5, 0.75, 1\}$

$\{\rho, \eta\} = \{1, 0.25\} \Rightarrow \beta = 2$

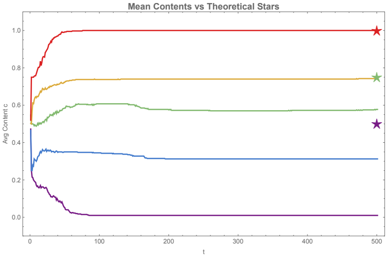


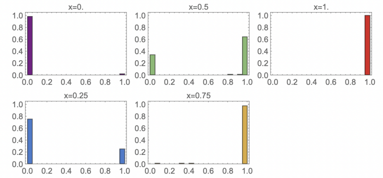




Commercial queries on new products with naive users

$\{\rho, \eta\} = \{0.6, 0.25\} \Rightarrow \beta = 2$





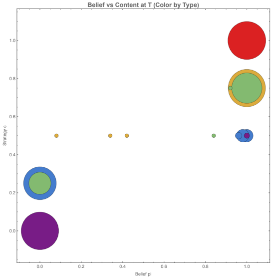
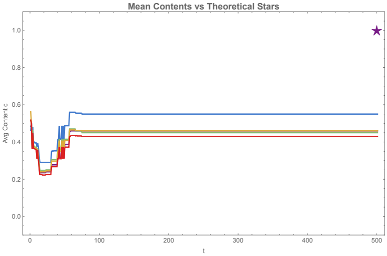
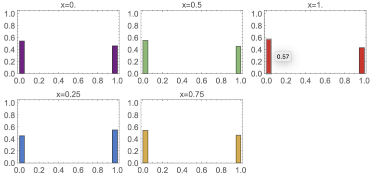


Figure: Simulations for $\omega = 1$

Pure informational queries from naive users

$\{\rho, \eta\} = \{0, 0.25\} \Rightarrow \beta = 2$





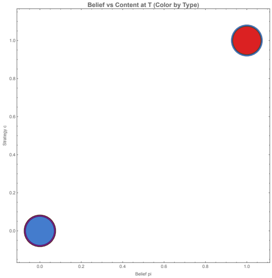
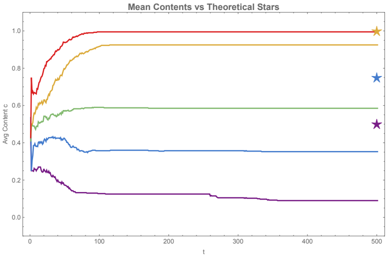
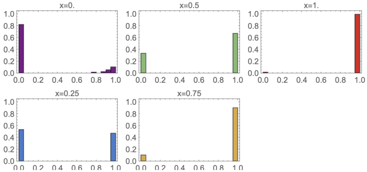


Figure: Simulations for $\omega = 1$

Information contested queries from naive users

$\{\rho, \eta\} = \{0.4, 0.25\} \Rightarrow \beta = 0.02$





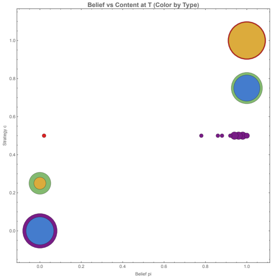


Figure: Simulations for $\omega = 1$

Endogenous distribution of firms' content

- At $\tau = -1$, Nature chooses $\omega \in \{0, 1\}$
- In every round $\tau \geq 0$:
  1. Each firm $j \in [0, 1]$ chooses the content $c_{\tau,j} \in X$ to be posted on its website
  2. $g_{\omega,\tau}(c)$: endogenous mass of firms posting content c at round τ given state ω
  3. The SE observes $c_{\tau,j}$ for all $j \in [0, 1]$
  4. The SE matches each user i with a firm j
  5. Users observe his/her personalized content and choose how much time to spend reading it

Website traffic

A firm j posting content $c_{\tau,j} = c$ obtains a *website traffic*

$$m_{\tau,j} = \underbrace{e^{-\beta\tau}}_{\text{traffic form SE exploration}} + \underbrace{(1 - e^{-\beta\tau})\frac{\hat{f}_{\omega,\tau}(c)}{g_{\omega,\tau}(c)}}_{\text{traffic from SE exploitation}}$$

where $\hat{f}_\omega(c)$ is the endogenous mass of user whose greedy content in round τ and state ω is $\hat{c}_{\tau,i} = c$

Firms' Q-learning algorithm

To choose its content $c_{\tau,j}$, each firm j uses a q-learning algorithm.

- With probability $e^{-\beta^F \tau}$, the firm's algorithm explores and sets $c_{\tau,j} = c$ with probability $g_\omega(c)$.
  Example: searching for information or using R&D output from the real world (reporters on the battlefield, R&D output, etc.).

- With probability $1 - e^{-\beta^F \tau}$, the firm's algorithm exploits and sets

  $$c_{\tau,j} = \arg\max_{c \in X} Q^F_{\tau,j}(c)$$

  where $Q^F_{\tau,j}(c)$ is the value that firm j's algorithm attaches, in round τ, to posting content c.

- The firm's algorithm aims to maximize traffic to the firm's website, but it can observe only its own traffic:

$$Q^F_{\tau+1,j}(c) = \begin{cases} \alpha^F m_{\tau,j} + (1-\alpha^F) Q^F_{\tau,j}(c) & \text{if } c = c_{\tau,j} \\ Q^F_{\tau,j}(c) & \text{if } c \neq c_{\tau,j} \end{cases} \quad (2)$$

Equilibrium definition

1. SE chooses α, β as to maximize average consulting time correctly anticipating firms and users behavior
2. Firms chose α^F, β^F as to maximize traffic to their website knowing SE's α, β and correctly anticipating users behavior
3. in each round τ
   - Users update their belief from observing $c_{,i}$ knowing $\alpha, \beta, \alpha^F, \beta^F$ and correctly anticipating $g_{\omega,\tau}$

$$\pi_{\tau,i} = \begin{cases} \frac{\pi_{\tau-1,i}g_{1,\tau}(c_{\tau,i})}{\pi_{\tau-1,i}g_{1,\tau}(c_{\tau,i})+(1-\pi_{\tau-1,i})g_{0,\tau}(c_{\tau,i})} & \text{if } c_{\tau,i} \neq \hat{c}_{\tau,i} \\ \frac{e^{-\beta\tau}(\pi_{\tau-1,i}g_{1,\tau}(c_{\tau,i})+1-e^{-\beta\tau})}{e^{-\beta\tau}\pi_{\tau-1,i}g_{1,\tau}(c_{\tau,i})+(1-\pi_{\tau-1,i})g_{0,\tau}(c_{\tau,i}))+1-e^{-\beta\tau}} & \text{if } c_{\tau,i} = \hat{c}_{\tau,i} \end{cases}$$

   - Users choose their consulting time $t_{\tau,i} = t^\star(x_i, c_{\tau,i}, \pi_{\tau,i})$

Results with endogenous contents

Results with endogenous contents

Preliminary conclusion

- Platforms' algorithms are extremely effective and lucrative for purely commercial queries and for consensual informational queries.
- For hybrid contested queries, machine learning-based search engines lead to inefficient outcomes, as users tend to be overly directed toward content that caters more to their type than to the best content given their type and the state of nature.
- % of users with $\pi_{500,i} < 1/2$ when $\omega = 1$:[1]

$\eta \setminus \rho$	0.0	0.2	0.4	0.6	0.8	1.0
0.25	23%	24%	23%	22%	29%	41%
0.50	12%	13%	11%	11%	14%	46%
0.75	4%	2%	5%	6%	10%	37%
1.00	5%	1%	3%	2%	9%	29%

- When users underestimate personalization:
  - extreme belief polarization emerges;
  - only purely commercial queries remain socially efficient;
  - consulting time (and thus, probably, the search engine's profits) exceeds the socially optimal level.

[1] $\rho = 0, \eta = 0.25, \beta = 0 \rightarrow 0.07\%$